

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 00:09:32 ; Search time 8498.8 Seconds

(without alignments)
31.610 Million cell updates/sec

Title: US-09-851-670-19

Sequence: 1 gctgcgtcgtgacccctcttgc 25

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 111874

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estopl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gq_estl:*
11: gq_est2:*
12: gq_hic:*
13: gq_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	68.0	42	13	A2938933 2M0197N10
2	14.6	58.4	43	13	A1115359 un84c06.r
3	14.2	56.8	30	13	A2310102 1M001BD24
4	14.2	56.8	32	10	BE317690 NF053E10L
5	14	56.0	55	11	BE977884 bs70d09.y
6	13.8	55.2	50	10	AU103198 AU103198
7	13.6	54.4	27	10	AM059815 LE7f08.y9
8	13.6	54.4	52	11	R42396 yf92a07.s1
9	13.6	54.4	57	10	A1019839 ua91h02.r
10	13.4	53.6	51	13	B05502 cSRL-64f6-u
11	13.4	53.6	60	10	AM004453 701931490
12	13.2	52.8	56	13	A2820546 2M0092207

13	13.2	52.8	58	10	AA415880
14	13.2	52.8	60	13	A2575892
15	13	52.0	32	13	A2306035
16	13	52.0	38	13	A2471063
17	13	52.0	39	13	A2428906
18	13	52.0	39	13	A2788235
19	13	52.0	42	11	T17509
20	12.8	51.2	36	13	A2807406
21	12.8	51.2	41	13	A2576441
22	12.8	51.2	52	11	BF643936
23	12.8	51.2	53	10	A1906734
24	12.8	51.2	55	13	A2755909
25	12.8	51.2	58	11	D38739
26	12.6	50.4	34	10	AA933070
27	12.6	50.4	46	10	A1357720
28	12.6	50.4	47	13	A2814250
29	12.6	50.4	51	10	AA672337
30	12.6	50.4	56	10	A1310538
31	12.6	50.4	58	10	A1340421
32	12.6	50.4	58	10	AA966109
33	12.6	50.4	60	11	BE857801
34	12.4	49.6	34	13	A2775556
35	12.4	49.6	39	13	A2610841
36	12.4	49.6	47	10	AA658865
37	12.4	49.6	49	10	A1567998
38	12.4	49.6	50	10	A1010205
39	12.4	49.6	50	10	A10107283
40	12.4	49.6	52	11	BF651215
41	12.4	49.6	54	13	A2612285
42	12.4	49.6	55	10	A1522415
43	12.4	49.6	60	11	BF638573
44	12.2	48.8	29	13	A0073365
45	12.2	48.8	32	10	A0014466

ALIGNMENTS

RESULT 1
A2938933 42 bp DNA GSS 26-APR-2001
LOCUS 2M0197N10R Mouse 10kb plasmid UNGC2M Library Mus musculus genomic
DEFINITION clone UNGC2M0197N10 R, DNA sequence.
ACCESSION A2938933.1 GI:13799132
VERSION A2938933.1
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE
AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tinger,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
Plasmid inserts
JOURNAL
COMMENT
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0197 row: N column: 10
Seq primer: CACACGAGAAACACGTATGACC
Class: plasmid ends
High quality sequence stop: 42.
Location/Qualifiers
1. 42

FEATURES
Source

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="MUGC2M0197N10"
/clone_l1b="Mouse 10Kb plasmid MUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (911473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

	Query Match	68.0%	Score 17:	DB 13:	Length 42;
	Best Local Similarity	80.0%;	Pred.	No. 2.5e+03;	
	Matches	20;	Conservative	0;	Mismatches 5; Indels 0;
Oy	1 gctgcgtcctgataccctcttcttc	25			
Db	41 ggtagaaccttgatcccatctttggc	17			

RESULT	2
A1115359	
LOCUS	A1115359 43 bp mRNA EST 02-SEP-1998
DEFINITION	uh84c06.r1 Soares mouse urogenital ridge NMR Mus musculus cDNA clone IMAGE:1764394 5' similar to SW-A472_RAT P53678 CLATHRIN COAT ASSEMBLY PROTEIN AP47 HOMOLOG 2 ;, mRNA sequence.

ACCESSION	AI115359	GI:3515683
VERSION	AI115359.1	
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE
AUTHORS
Euarctota; Metazoa; Chordata; Crustacea; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 43)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.

TITLE	The Washu-HHMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Marra M/Mouse EST Project

Email: mouseest@atson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:963918
Trace considered overall poor quality
Possible reversed clone. Similarly on wrong strand
Seq primer: -28m13 rev2 ET from Amersham

FEATURES		SOURCE	
High quality sequence stop: 1.			
Location/Qualifiers		1..43	
/organism="Mus musculus"			
/db_xref="taxon:10090"			
/clone_image="IMAGE:1764394"			
/clone_lib="Soares mouse urogenital ridge NKUR"			
/sex="equal ratio of male:female"			
/tissue_type="urogenital ridge (embryonic)"			
/dev_stage="fetal, mixture of 11.5 and 12.5 dpc"			
/lab_host="DH10B"			
<p>/note="Organ: gonad; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15' TGTTACCAATCTGACAGCGAGCGCGCATTTCTTTTTTTTTTTTTTTTTTTT T3'); (double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fátima Bonaldo."</p>			
BASE COUNT	5 a	11 c	7 g 14 t 6 others
ORIGIN			

Query Match	58.4%	Score 14.6	DB 10	Length 43
Best Local Similarity	77.3%	Pred. No. 2.5e+04		
Matches 17	Conservative	0	Mismatches 5	Indels 0
			Gaps 0	
Qy	1	gctgcatctgfatcccttctt	22	
Db	12	GCTCTNNTGTCGATTA	CTCTT	33

RESULT	3
AZ310102	
LOCUS	AZ310102 30 bp DNA GSS 29-SEP-2000
DEFINITION	IHM018D2R Mouse 10k plasmid UGCG1M library Mus musculus genomic
ACCESSION	clone UGCG1M0018D24 R, DNA sequence.
VERSION	AZ310102
	AZ310102.1 GI:10351755

SOURCE	ORGANISM
house mouse.	<i>Mus musculus</i>
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus (bases 1 to 30)
REFERENCE	Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamli, C., et al.

TITLE	Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL	plasmid inserts
COMMENT	Unpublished (2000)
Contact:	Robert B. Weiss

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunne@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0018 row: D column: 24
 Seq primer: CACACAGCAACACGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 30.
 Location/Qualifiers
 1. .30

```
1.30
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U08C1M0018D24"
/clone_1lb="Mouse 10kb plasmid U08C1M library"
```

```

/sex="Male"
/lab-host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42ny, Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g1147321149b/AP129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT      2 a      14 c      5 g      9 t
ORIGIN

Query Match      56.8%; Score 14.2; DB 13; Length 30;
Best Local Similarity 84.2%; Pred. No. 3.4e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy      2 ctgcactgtatccctc 20
      ||||| ||||| |||||
Db      12 CTCGCTGTCTGACGCTTC 30

RESULT 4
BE317690      52 bp      mRNA      EST      21-DEC-2000
LOCUS      NF053E10LF.F1071 Developing leaf Medicago truncatula cDNA clone
DEFINITION      NF053E10LF 5', mRNA sequence.
ACCESSION      BE317690
VERSION      BE317690
KEYWORDS      GI:11960827
SOURCE      EST.
ORGANISM      Medicago truncatula
      barrel medic.
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
      Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
      Medicago.
      1 (bases 1 to 52)
      Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
      Flores, H.R., Imman, J.T., Weller, J.W. and May, G.D.
      Expressed Sequence Tags from the Samuel Roberts Noble Foundation
      Medicago truncatula leaf library
      Unpublished (2000)
      On Jul 14, 2000 this sequence version replaced gi:9191563.
      Contact: May GD
      Plant Biology Division
      The Samuel Roberts Noble Foundation
      210 Sam Noble Parkway, Ardmore, OK 73402, USA
      Tel: 580 221 7391
      Fax: 580 221 7380
      Email: gdmay@noble.org
      Medicago Genome Initiative accession: MGI:S:21873
      Insert Length: 776 Std Error: 0.00
      Plate: 053 row: E column: 10
      Seq primer: TCACACGAGAAACAGCTATGAC.
      Location/Qualifiers
      1..52
      /organism="Medicago truncatula"
      /db_xref="taxon:3880"
      /clone="NF053E10LF"
      /clone_lib="Developing leaf"
      /tissue_type="leaf"

```

```

/dev_stage="pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of very
young, developing, mature and senescing leaves."

BASE COUNT      11 a      16 c      14 g      11 t
ORIGIN

Query Match      56.8%; Score 14.2; DB 10; Length 52;
Best Local Similarity 84.2%; Pred. No. 3.7e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy      2 ctgcactgtatccctc 20
      ||||| ||||| |||||
Db      38 CTCGAGGTGTGATACCTTC 20

RESULT 5
BE977884      55 bp      mRNA      EST      04-OCT-2000
LOCUS      bs70809.y1 Drosophila melanogaster adult testis library Drosophila
DEFINITION      melanogaster cDNA clone bs70809 5', mRNA sequence.
ACCESSION      BE977884
VERSION      BE977884
KEYWORDS      GI:10608804
SOURCE      EST.
ORGANISM      Drosophila melanogaster
      fruit fly.
      Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
      Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
      Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
      1 (bases 1 to 55)
      Andrews, J., Bouffard, G. and Oliver, B.
      Drosophila melanogaster testis expressed sequence tags
      Unpublished (1999)
      Contact: Brian Oliver
      Laboratory of Cellular and Developmental Biology
      NIDDK, National Institutes of Health
      6 Center Drive MSC 2715, Bldg 6, Rm B1-13, Bethesda, MD 20892 USA
      Fax: (301) 496 5239
      Email: oliver@helix.nih.gov,
      http://www.nidck.nih.gov/intram/people/boliver.htm
      Tissue isolation and library construction performed at the National
      Institute of Diabetes and Digestive and Kidney Diseases, NIH (see
      http://www.nidck.nih.gov/intram/people/boliver.htm). DNA sequencing
      and analyses performed by National Institutes of Health Intramural
      Sequencing Center (NISC; see http://www.nisc.nih.gov).
      Plate: 70 row: d column: 09
      Seq primer: M13RPI reverse primer (ABI).
      Location/Qualifiers
      1..55
      /organism="Drosophila melanogaster"
      /strain="y[*] w[67c1]y"
      /db_xref="taxon:7227"
      /clone="bs70809"
      /clone_lib="Drosophila melanogaster adult testis library"
      /sex="male"
      /dev_stage="1-5 day adult"
      /lab_host="SOPK (Stratagene)"
      /note="Organ: testis; Vector: pBluescript SK (Stratagene);
      Site_1: EcoR I; Site_2: Xho I; Testes dissected from 1-5
      day adult y[*] w[67c1]Y males raised at 25°C. RNA
      isolated using Trizol (Life Technologies) and a single
      round of poly(A)+ selection using Oligotex (Qiagen). cDNA
      library constructed using Stratagene Zap-cDNA synthesis
      kit. Oligo dt-primed, size fractionated -1-6 kb, and
      directionally cloned at EcoRI and XhoI in uni-ZAP XR.
      Following a single round of amplification pBluescript SK
      phagemids were mass excised. A distribution channel for
      clones is being sought, but not currently available.
      Requests for clones cannot be honored."
      8 a      11 c      19 g      17 t

```

Query Match 56.0%; Score 14; DB 11; Length 55;
 Best Local Similarity 77.3%; Pred. No. 4.5e+04;
 Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 cgaatctgataccctcttctgc 25
 ||| | | | | | | | | |
 Db 4 CGAGCGATCATGCTCTCTTGC 25

RESULT 6

AM03198

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match 55.2%; Score 13.8; DB 10; Length 50;
 Best Local Similarity 72.0%; Pred. No. 5.4e+04;
 Matches 18; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 gctgactgataccctcttctgc 25
 ||| | | | | | | | | |
 Db 20 GATGATCTGTGCTCTCTCTCGC 44

RESULT 7

AM059815

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 27)
 Brenner, S., Williams, S.R., Vermass, E.H., Storck, T., Moon, K.,
 McCallum, C., Mao, J.I., Kirchner, J.D., Elert, S., Dubridge, R.B.,
 Burcham, T. and Albrecht, G.
 In vitro cloning of complex mixtures of DNA on microbeads: Physical
 separation of differentially expressed cDNAs
 Proc. Natl. Acad. Sci. U.S.A. 97 (4), 1665-1670 (2000)

MEDLINE 20144098
 COMMENT
 Contact: Burcham TS
 LYNX Therapeutics, Inc.
 25861 Industrial Blvd., Hayward, CA 94545, USA
 Tel: 510 670 9338
 Fax: 510 670 9302
 Email: timbelynxgen.com
 Sequence obtained from LYNX Therapeutics Megasort technology.
 Collected from the down-regulated gate.
 High quality sequence stop: 27.

FEATURES

source

BASE COUNT

ORIGIN

Query Match 54.4%; Score 13.6; DB 10; Length 27;
 Best Local Similarity 80.0%; Pred. No. 6e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 6 atctgataccctcttctgc 25
 ||| | | | | | | | | |
 Db 8 AGCTGATCTCTCTCTCTC 27

RESULT 8

RA2396/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 52)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston
 R., Williamson, A., Wohlmann, P. and Wilson, R.
 The Washu-Merck EST Project
 Unpublished (1995)
 Contact: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 Insert Size: 1268
 High quality sequence starts: 1 High quality sequence stops: 1
 Source: IMAGE Consortium, LNU. This clone is available royalty-free
 through LNU; contact the IMAGE Consortium (info@image.lnu.gov)
 for further information. Trace considered overall poor quality
 Insert Length: 1268 Std Error: 0.00
 Seq Primer: Promega -21m13
 High quality sequence stop: 1.
 Location/Qualifiers
 1..52
 /organism="Homo sapiens"
 /db_xref="GDB:402299"
 /db_xref="taxon:9606"

FEATURES

source

```

/clone="IMAGE:29952"
/clone_lib="Soares infant brain INIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone="IMAGE:1364883"
/clone_lib="Soares_mammary_gland_NBMNG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT773D-Pac (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the lacmid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."
I - oligo(dT) primer [5];
AAGTGGAGAAATTCGCGCCGAGAAATTTTCTTTTCTTTT 3'};
BASE COUNT      11 a      12 c      14 g      12 t      3 others
ORIGIN

Query Match      54.4%; Score 13.6; DB 11; Length 52;
Best Local Similarity 76.2%; Pred. No. 6.5e+04;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 5 gacgtgtgacccctcttctg 25
    ||| ||| ||| ||| ||| |||
Db 37 GANCTGAGATCAGTCATTC 17

RESULT 9
LOCUS      A1019839      57 bp      mRNA      EST      16-JUN-1998
DEFINITION ua91h02.r1 Soares_mammary_gland_NBMNG Mus musculus cDNA clone
IMAGE:1364883 5' similar to gb:Y00516 Mouse mRNA for aldolase A
(MOUSE); mRNA sequence.
A1019839
A1019839.1 GI:3234175
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 57)
AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
            The WashU-HMI Mouse EST Project
            Unpublished (1996)
            Contact: Marra M/Mouse EST Project
            WashU-HMI Mouse EST Project
            Washington University School of Medicine
            444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:898103
            Seq primer: -28m13 rev2 ET from Amersham.

FEATURES
Source
Location/Qualifiers
1..57
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1364883"
/clone_lib="Soares_mammary_gland_NBMNG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT773D-Pac (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)

```

```

primer [5];
TGTTRACCAATCTGAGTCGGAGCGCCGAGATGTTTTTTTTTTTTTTTTTTTTT
T 3'}; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."
18 a      17 c      8 g      14 t
BASE COUNT
ORIGIN

Query Match      54.4%; Score 13.6; DB 10; Length 57;
Best Local Similarity 80.0%; Pred. No. 6.6e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 gctgacgtgtgaccccttc 20
    ||| ||| ||| ||| ||| |||
Db 6 GCTAATCTGCCATCCCTTC 25

RESULT 10
LOCUS      B05502      51 bp      DNA      GSS      13-JUL-1996
DEFINITION CSRL-64f6-v CSRL flow sorted Chromosome 11 specific cosmid Homo
sapiens genomic clone CSRL-64f6, DNA sequence.
B05502
B05502.1 GI:1414780
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 51)
AUTHORS    Evans,G.A., Burbee,D., Davies,C., Hahner,L., Oliver,T., Gilbert,M.,
            Jones,D., Ward,T., Gillilan,E., Schagemann,J., Probst,S., Harris
            J., Bedford,J., McFarland,J., Buzlinski,K., Khan,M., Kupfer,K. and
            Garner,H.R.
            Genomic Sequence Sampled Map of Chromosome 11
            Unpublished (1996)
            Contact: Evans GA, Shane Probst
            McDermott Center for Human Growth and Development
            University of Texas Southwestern Medical Center At Dallas
            5323 Harry Hines Blvd, Dallas TX 75235-8591
            Tel: 214-648-1600
            Fax: 214-648-1666
            Email: geyvans@utsw.swmed.edu, shanemcdermott.swmed.edu
            Seq primer: T7
            Class: cosmid ends
            High quality sequence stop: 51.

FEATURES
Source
Location/Qualifiers
1..51
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSRL-64f6"
/clone_lib="CSRL flow sorted Chromosome 11 specific
cosmid"
/sex="female"
/cell_type="chimeric hamster somatic cell hybrid"
/note="Vector: scos-1; Human chromosome 11 specific cosmid
library prepared from flow sorted human Chromosome 11
derived from Chinese Hamster Ovary (CHO) monochromosomal
somatic cell hybrid, J1"
BASE COUNT      2 a      11 c      16 g      21 t      1 others
ORIGIN

Query Match      53.6%; Score 13.4; DB 13; Length 51;
Best Local Similarity 70.8%; Pred. No. 7.9e+04;
Matches 17; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 gctgacgtgtgaccccttc 24
    ||| ||| ||| ||| ||| |||

```


JOURNAL
COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
WCI:477804

FEATURES
source

Seq primer: -40m13 fwd. ET from Amersham.

Location/Qualifiers

1. 58
/organism="Mus musculus"
/strain="B6D2 F1/J"
/db_xref="taxon:10090"
/clone_id="IMAGE:793396"
/clone_lib="Knowles Solter mouse 2 cell"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"
/note="Organ: embryo; Vector: pBluescribe (modified);
Site_1: Mu1; Site_2: SalI; Cloned unidirectionally from
mRNA prepared from 13,500 2-cell stage embryos. Primer:
SalI(dT): 5'-CGGTGACGCGCGACGCTTTTCTTTT-3', CDNAS
were cloned into the Mu1/SalI sites of a modified
pBluescribe vector using commercial linkers (NEB).
Average insert size: 1.2 Kb."

BASE COUNT 19 a 10 c 27 t

ORIGIN

Query Match 52.8%; Score 13.2; DB 10; Length 58;
Best Local Similarity 83.3%; Pred. No. 9.6e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 atcgtgacgtccctt 23
||||| ||||| |||||
Db 25 ATCTCATTCCTCTTT 42

RESULT 14
A2575892 60 bp DNA GSS 06-DEC-2000
LOCUS
DEFINITION AST-T23B0019 Genetrap T47D Human Breast Carcinoma Library Homo
sapiens genomic 5', DNA sequence.
VERSION A2575892
KEYWORDS GSS.
SOURCE
ORGANISM human.

REFERENCE
AUTHORS Homo sapiens
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL 1 (bases 1 to 60)
COMMENT 1 Henkel,G., Livanage,M., Pratt,E., Huang,D., Riley,M., Bernardino,A.
Exon-trap tags from a T47D GenomesScreen(TM) Library
Unpublished (2000)
Contact: Greg Henkel
Gene Expression
Aurora Biosciences Corp.
11010 Torreyana Road, San Diego, CA 92121, USA
Tel: 8584048436
Fax: 8584046719
Email: henkel@aurorabio.com

Pools of cells were isolated from a GenomesScreen(TM) library. The
library of cells was generated by retroviral integration of a gene
tagging element consisting of: 1) A promoterless beta-lactamase
preceded by a splice acceptor as a reporter for gene expression;
2) A promoter driving neomycin resistance followed by a splice
donor to trap downstream exons. 3' RACE from neomycin gene was
performed using total RNA from isolated pools. Output was shotgun

FEATURES
source

cloned in pAMP-1 and used to transform DH5-alpha competent
bacteria. 5' ends of reported sequences were immediately preceded
by splice donor from the trapping construct.
Class: exon-trapped.

Location/Qualifiers

1. 60
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="Genetrap T47D Human Breast Carcinoma Library"
/tissue_type="Carcinoma"
/cell_type="Epithelial"
/cell_line="T47D"
/note="Organ: Breast; Vector: pAMP-1; 3' RACE of total RNA
from genetrap pools; shotgun clone in pAMP-1 and used to
transform DH5-alpha competent bacteria."

BASE COUNT 14 a 16 c 16 g 14 t

ORIGIN

Query Match 52.8%; Score 13.2; DB 13; Length 60;
Best Local Similarity 83.3%; Pred. No. 9.7e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 atcgtgacgtccctt 19
||||| ||||| |||||
Db 8 CTCATCTGTCCACCAT 25

RESULT 15
A2306035 32 bp DNA GSS 29-SEP-2000
LOCUS

DEFINITION 1M0006M23R Mouse 10kb plasmid UGECIM library Mus musculus genomic
clone UGECIM0006M23 R, DNA sequence.
VERSION A2306035
KEYWORDS GSS.
SOURCE
ORGANISM house mouse.

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scturognathi; Muridae; Murinae; Mus.
1 (bases 1 to 32)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0006 row: M column: 23
Seq primer: CACACAGGAACAGCATGTGACC
Class: plasmid ends
High quality sequence stop: 32.

JOURNAL
COMMENTFEATURES
source

Location/Qualifiers

1. 32
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_id="UGECIM0006M23"
/clone_lib="Mouse 10kb plasmid UGECIM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD29v. Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adsorbed DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (q147321419b1aFN12072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

BASE COUNT	6 a	12 c	3 g	11 t
ORIGIN				

ORIGIN

Query Match

52.08; Score 13; DB 13; Length 32;

Best Local Similarity 76.28; Pred. No. 1.1e+05;

Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ctcgatactgtgatacccttctt 22

— — — — —

Db 12 CCCCATCTGTAAATGCCCTCTT 32

Search completed: March 9, 2002, 00:09:35
Job time: 11051 sec

Job time: 11051 sec